From Data Mess to Data Mesh: Navigating People, Process & Platforms • Nicki Watt • GOTO 2022 - YouTube

https://www.youtube.com/watch?v=TONOUCAkfLE

(21:44) different the way that the principles are implemented will be quite different depending on on sort of where you're coming from so okay with that sort of framework in mind for considering the options laid out to make this a little bit more practical let's look at some sort of real sort of problems that we see clients coming to us with and what they present with and see how how this works if we go through this process so first up what kind of problems or what kind of mess do we see clients presenting with

(22:16) so they will often come in various sort of shapes and sizes and they can just be described in different ways by different groups so if we look at this from the perspective of maybe a data scientist or an analyst they will come and they'll say things like I'm not able to find my data easily there's too many systems that are all over the show and you know I just I simply can't find it I don't know where to start or and this is the more common one they'll come and they say well actually

(22:42) my starting point is a is a data Lake and so I I just can't easily access or make sense of the data in there there's just this dump of files and things and as a an analyst I'm really struggling to be able to work out what's in there and and use it and that's because with a data Lake architecture the files are usually dumped and the onus is on the consumer to be able to make sense of it so other angles you might look at this from the sort of ctOS perspective and the problem statements that we we see

(23:15) coming up here is really around well the projects are taking too long um and uh you know the business is losing out on opportunities and whilst this may apply to General sort of operational projects in this case it's usually related to the fact that they've got a project where you've got some data-driven insight that they're trying to use in order to sort of reduce the cost of the business or increase the revenue and it's those type of projects that are really struggling and this is closely related to one of

(23:44) the other things we see coming up quite often and that is that teams are overloaded and this is generally um I mean it's General problem in in many cases but not having enough people to to do the work is is something that we see quite often and specifically bottlenecks when it comes to data teams and centralized data teams so they might also have uh sort of challenges on the security or governance front worried about things like potentially exposing unnecessary sort of pii data etc etc so how can we go about solving these

(24:17) these challenges so let's focus in on the on the data scientists problems and and go through our process and see what we come up with now please note like normally you don't just take this one thing and and go through this process you'd usually do it across um the whole range of things to get a holistic view across across everything but for the purposes of this talk we're just going to go through that and the reason why it's interesting to go through the data data scientists or the analysts sort of problems is that

(24:47) like I say it can be raised at different levels so um if if your problem's raised by a C Level or that's uh an executive or something you can generally be certain that it's of strategic value but if it comes through the the sort of I.T Department it may just be a local challenge but it may actually be something that that is strategically important but you need to dig a little bit deeper to actually find out if that is the case so we have some problem that's been raised we've established that they're

(25:18) having difficulty in finding data and or making sense of it we understand uh you know what's needed they need this to sort of build and train models and extract insights Etc so as we've seen before rather than just assuming that data mesh is the right fit we start asking questions you know why is it important to actually address this and you might get answers like well if I can more easily find and address my data then it's easier and faster for me to build my models and that's a valid answer

(25:50) except we keep asking more it's not the real reason behind it so we say well why why is that important and they'll say well the models are really the key part to how we deliver insights and value to our users and we're growing and you know we can't afford to sort of slow down in this area I'm under a lot of pressure as the analyst because we keep getting requests to to sort of do the work and and we can't really kind of keep up that's still we keep going deeper sometimes you might get an answer like

(26:18) well my boss has said I need to solve this so that's why I'm here but not relevant you just keep going until hopefully you get to some of the more sort of fundamental reasons for uh for why this is happening and in this case it may be that well actually the competition is is already doing this and um they're already using AI models and things to reduce their costs or increase revenue and there's an acknowledgment that actually if if we don't do this we're going to get left behind as a

(26:44) business now we've really hit the the sort of nail on the head and the I found that sort of the business driver behind it so that's great uh We've we've hopefully established what the what the sort of problem is and the business driver behind it so we'll Replay that back uh to to the client and you know just make sure that they they agree with the statement uh so we'll say they're having difficulty but the main reason is really for the business growth um the business is growing the

(27:13) competition is increasing and say yes I I agree with that so now we move into all right well how do we actually solve this and this is where there are going to be a variety of different ways to solve this there's many ways to sort of do this and there's a few listed here and I'm not saying these are the only ones I'm just using it for illustration purposes um but maybe you know can we just do this simply with with technology why why can't we just you know sort of implement a data catalog and you know in one ways this is this is

(27:45) quite sort of simple and it will address a decent proportion of the of the problem that was raised um you know providing a single place to have your sort of data sets registered and searched and found will answer some of that question but um uh you know even the seemingly technical R sort of approach will require still a decent amount of having to run around and find out you know whether who owns the data etc etc but um it's not going to do everything for us our challenges are a bit more sort of strategic and we also one of the

(28:19) things that was raised was that we need to be able to easily access the data so this is part of the solution but it's maybe not the whole the whole thing then we say okay well what about uh what about a data lake house so um this might be a strategic and a sort of technical approach and for those of you that maybe aren't to wear a data lake house is generally where you try to combine The Best of Both Worlds of a data warehouse and a data Lake to to bring about an easier way to make things accessible to uh to

(28:49) to analysts using it give them some sort of structure around it but also still allow for the mass ingestion of of stuff and whilst data is maybe currently unusable in its existing sort of forming a data like the promise as I say is to combine these two and we might say well actually this might be quite close like a data a data lake house could be a good could be a good answer for us and you should actually look at whether that's whether that fits for you but we're looking at the data data mesh side of things so you said all right

(29:23) let's let's see how this fits against the objectives that we're we're trying to look at doing so um as I said before uh if we have a look at what some of the objectives of data mesh were um in the beginning let's see how closely we actually align with that with our example so do we really have a complex um and fast changing setup so as a business you might say well yes we're we're doing a lot of M A Acquisitions we've got lots of new business units popping up so it's you know there's a

(29:55) lot of complexity in in our system so yes probably tick that box are we growing fast yes uh the according to the sort of the problem statements that we saw earlier that seemed to be the case are we really a data driven organization are we putting Ai and insights at the heart of what we're trying to do and it sounds like in this case they really are because you know they're trying to reduce costs and increase Revenue Etc through the through the things that they're doing so it seems like a good fit so all the

(30:26) lines so hopefully data mesh is is a valid option at this point um and you know we can we can go forward but just because there's a theoretical alignment here doesn't mean that it's going to work out and we now ask the really crucial question which is how far are you prepared to go in order to do this now this is the stage that I think scuppers most organizations that we see um uh when we go on these Endeavors and I really love this quote um which says you know everybody wants to go to heaven but nobody wants to die and I I heard

(31:01) it's actually in a previous talk about fast flow in a government in in sort of Norway um but whilst I find this profound and deep on many levels it sums up the essence I think of what is needed to really do a data mesh well and that is you know whilst the gains and the sort of outcomes of a data mesh may sound Heavenly to some people are you really prepared to do what it takes to get there and this will often involve giving up some preconceived ideas often about centralized control and also changing and letting go of some of the processes

(31:33) and ways of working that you may hold dear so what might be impacted uh what kind of things might you have to look out for here so there's potentially changes to your team structures your operating models governance architecture so quite often moving down this path you need to move from having project-based teams to product based teams so this is quite a change to a product mindset and this is a this is something that would need to change in in the way that the organization kind of works also if we think about the fact that

(32:06) we've got these domains who are now going to be responsible for the data and exposing it there's a new set of type of roles which get introduced into into the organization so you have something called a data product owner so who's going to be that data product owner are they really going to be able to be responsible for the the data products that are being put out there are they incentivized to do it because you know some argument people might say well you know if I have to put data out there but

(32:34) I'm not using it why why should I bother so how do you go about actually incentivizing people to do that and it's often a lot easier said in theory than it is in practice on the governance side of things are you really prepared to allow the individual teams to do their own governance so if you're an organization that's really doing a lot of command and control in the governance section this is really difficult for many organizations and architecturally the cherished command and control like I

(33:01) say moves to more decentralized approach but if you're there great we go down that path if you're not it's not a problem you look at the other options work out what what might be appropriate for you and maybe maybe you're only halfway there there's some political pressure or deadlines and you can only go so far that's okay you can you can look at other things whatever you're doing uh irrespective of that you always want to check am I improving Against The Core Business objectives that I I set out to do so

(33:31) here's an example if you're in a sort of an HR domain uh with these problems you might say well I was expecting to you know I'm ingesting a whole bunch of CVS or something and I want to make sure I have a reduction in my the candidates in the first phase of hiring or whatever they may be you use this to actually work out am I making progress or not so having gone through a sort of practical uh sort of walkthrough of what that might look like with some problems we've seen that data mesh is not necessarily the answer

(34:06) for everybody but even if it is even if it fits the or it takes all the boxes for the theoretical reasons why I should do it you really do need to make sure that you're prepared for the the potential changes that come with that and it can require big changes to to organizations and the ways of working and this is this is really the the area that we see companies struggling with if they don't recognize this this early on because sometimes they might start this and then realize oh this this is really a lot more than I was expecting in the

(34:38) first place so but what if you can't fully Implement all of the aspects what if you can say well I I could do some of it but maybe not not all of it does this preclude you from from doing data mesh is it an All or Nothing approach so in this section um also try and give you some sort of tips and tricks that we've kind of seen as we as we've worked with clients on this so in all cases we always recommend to people that you start small so don't try and boil the whole ocean you find one or two sort of use cases that really kind

(35:14) of make sense preferably not on the critical path that you can Implement and you can try and make progress along each of these towards each of these principles now you know whether you're going data mesh or not you can still take actions uh in terms of moving towards these these areas with outgoing necessarily the whole way so first up whether you're doing data mesh or not I'm going for a cloud first approach makes everything a whole lot easier so this is my first sort of top tip that we've that we come across and that we

(35:50) recommend to clients is always to lay a solid foundation by starting with the cloud so the cloud has got this sort of near infinite scaling capabilities and if you are an organization that itself is really scaling quite fast it's something that's obviously going to be really important to making sure that that's part and parcel of it it's got modern data tools and stacks on it and the thing that is is really helpful is that most of the clouds out there have got something called like a well architected framework so this

(36:18) provides you with a sort of a template for the the way that you can build some of your architectures in the clouds to make sure that you can bring together all of the right services to kind of do things that you need but also to exhibit some of the desired characteristics that most sort of projects and and solutions are going to need so things like being able to do self-service being able to have your scalability and security this will come for free in a way if you actually follow the well-architected guides so this is as I say this is the

(36:52) Azure example which talks about Landing zones but Amazon has got a a similar thing and if you if you go into if you go into these sort of reference architectures you'll see they've actually got a section now for data mesh so as more and more people are implementing their Solutions and we're learning a bit more they'll say well you know if you're in this circumstance you might want to look at doing it like this or these are a few options so I do recommend if you're starting out to start start with a well

(37:19) architected framework you obviously have to customize it but it's a good starting point so as I said before if your starting point uh maybe was a data warehouse or you've just got a data Lake what we see as one of the alternate sort of options is is either a data lake house architecture and in some cases I would say that's actually not necessarily a bad step on the path but you have to be quite careful about how you're doing it if you ultimately want to go to the data mesh solution so some people see that the lake house

(37:56) architecture as a distinct strategy in and of itself and like I say it may be but if it's done with care it can really form the underpinnings of a data mesh architecture for you as well so I'm going to explore this a little bit by taking a little bit of a detour by investigating the concept of like how much do you centralize how much do you decentralize so my second sort of top tip here is really to be pragmatic about the level of decentralization that you have to go to when you're implementing this so

(38:26) centralization is often seen as a dirty word in data mesh because it's all about decentralizing things but you want to do everything in context and don't throw the baby out with the bath water so there are obviously benefits to decentralization you've got things like you know better ownership there's less contention it's easier to scale and many organizations will naturally they'll they'll they'll be quite happy with that and they may have already gone through this as part of their digital

(38:54) Transformations and things like that however not everybody's convinced about um at this at a technical level so some of the questions which then kind of come back are saying well what about the cloud costs you know if we've got if all the domain teams are doing their own things and and stuff what about the egress costs that are going to happen or is my is my cloud ball going to suddenly sort of massively spin out or if everybody has to spin up all of this infrastructure that's surely going to cost a lot more money

(39:24) and there's worries about data silos and performance of having to bring all of these data products together and how do we go about solving this well we ask some more questions so things like what is actually the most important thing to you so is the most important thing the agility that you get by being able to scale quickly so if I bring a new business unit on board I can just spin up my self-contained sort of team and off I go and and that is really being able to respond quickly is is more important than maybe the costs or other

(39:56) kind of factors but if it's you know if it's cost if you're cost sensitive and or other kind of factors you do you know you have to kind of take that into account and the way to do this is is with through compromise basically so if we have a if we have a look here we can see that in a really sort of traditional sort of um centralized data Lake world it's very common to have Central uh data teams so you might have um well Central teams maybe Central Storage as well as sort of platform capabilities

(40:29) but what we find is that the problem is not necessarily the storage sort of element itself but often it's the the contention around the people and the functions and the time needed to actually do the work so for example you might find that if you if you see you've got a central data team a warehouse team sorry an ETL team or something like that there's contention because you're trying to get the data out of the operational stores that are in the project teams and make it available for the analysts and

(40:57) there's only a certain number of people and that's really where we see a lot of the The bottlenecks Happening which is at the sort of organizational level now there can be contention obviously on the infrastructure and that side of things but um this is more this is more typical of of actually what we see now at the Other Extreme you have a fully decentralized data mesh setup and this will include you've got fully self-contained domain aligned teams all looking after the data themselves producing their data as a product that

(41:29) everybody can use it using the self-service platform and all of the data is stored locally within that within that domain now this is fully scalable but it is quite complex and it comes with challenges and like I say it is sometimes too much for for some people so what do we do you know there's usually a compromise of sorts that that we can have here and it usually involves some aspects of having the platform move a little bit more towards the sort of centralized side of things but at the same time holding true to the the

(42:02) broader decentralization concept of having decentralized an operating model and and teams on top of that now whilst is the physical architecture itself May Implement you know some sort of centralization in the storage layer and and this is maybe where things like a a lake house architecture do actually come into play so some people may say as I said they might say well this is not really a data mesh because you've got some kind of sort of centralized storage and stuff like that but I would say actually it

(42:34) still adheres to the principles of data mesh so for example you've still got the um the domain is responsible for the data uh in in their area granted it might be on a shared platform but the the platform team makes it easy for the individual teams to control and manage that portion of the of the of the sort of area that they're working with so they can still control their own destiny it's just that there's there's some kind of sharing of of resources and things to help things with like costs and stuff

(43:05) like that so data mesh is a concept as it's described in in xamax book is very silent on how to do this at an implementation level and as I say we we kind of see these compromises happening so in this case you kind of remove some of the contention that happened at the the integration team level um and you know whilst you still have to be aware of the fact that there might be physical contention you have helped to like I say maybe deal with some of the the costings by by using shared platforms and stuff like that

(43:38) now this is sort of meshy it may not be perfect but as I say we're all kind of still learning and trying to work out what actually makes sense now in reality as I say you're going to be in you know different companies are going to be in quite different sort of setups and there's usually a you know a variety of things that you can do here so you might have different topologies and you might have a hybrid case where you've got some stuff that's isolated some that's shared and yet more still

(44:04) we've actually seen when you go into really big businesses there is you know if you've got individual lines of of business they may actually duplicate this whole stack and and kind of do things that way so multiple different ways to do this is this still a data mesh maybe maybe not does it matter um maybe time will tell but for me the crucial question is really doesn't matter what your you know how you've kind of done it are you really are you still improving are you actually addressing the original

(44:36) challenges that were raised and we again we kind of look at the kpis and and see whether we're making progress or not now there's other accesses that we can improve along so and by the way I'm not saying you have to do this all in this order I'm just using this as an example but in terms of maybe moving more towards the domain orientated ownership you may say well at a minimum I want to be able to find and assign owners to to data now even if you haven't split as an organization into sort of domains yet

(45:07) this is a small tap a small step that you can take on the journey to at least find out like where your data is and who the data owners are and it helps you to actually work out maybe where some of those boundaries and Things Are maybe next you might say okay well on top of that I'll establish a data set a sort of catalog just kind of make sure where everything is it's maybe not a full-on data product catalog yet but but it's still it's it's you know it's a good start but when you start getting towards

(45:39) um more sort of serious about actually I really I really do want to go down the the data mesh side of things this is where all of the operating models and that whole side of things needs to change so it may be that at this stage you say well actually maybe I need to pause because there's a fundamental change which needs to happen in our operating model so maybe there's some kind of digital transformation that that we need to kind of address or there may even be a digital transformation or something like that going on and you

(46:07) need to hook into that um so again like I say you need the opportunities to shape what what that's going to look like to ensure that the different roles that are needed to do data mesh well can actually be taken into account this is also involved sort of moving from Project teams to product teams and and things like that so my top tips here is really to sort of start small learn adapt rinse repeat so something that we find is what I'd say is the data mesh killer is if you're trying to do this and you go to the

(46:40) organization say well you have to do a full-on reorg right from the beginning and everybody needs to be split up into domain teams right from day one that generally doesn't work and the way to go about doing this is more to as I say find some use cases that really make sense do one or two domains and then try and sort of iterate around that and learn learn from that before you you kind of scale out into into the other areas so you can piggyback off a existing sort of transformation if that's going on

(47:11) this is a good opportunity to do that and something that we found actually really quite helpful in first of all finding the domain boundaries but also defining your data products like what exactly is a data product what is in it what does it look like is to use something called event storming and this we have found really a really helpful technique in order to do that so you know you carry on you maybe do a data product catalog and as you can see we're kind of moving a little bit further along all of the sort of

(47:43) accesses here you might extend that catalog now to have pii data classifications etc etc um but you know as we're kind of going on we're continually learning we are evolving and if you recall from the illustration that I have before everything is connected and interacting so as you do some of these things you might find I now need to go and refactor my platform because now I've changed the way my my teams work and are structured and I need to do some sort of changes here or there's various bits that you might need

(48:12) to change along the way um but the point is that you can make progress without doing the whole thing as long as you start small and you kind of iterate and Implement sort of small small things so the other thing the other top tip that we found however in doing this is that um the tooling is really really immature in this space so you're going to have to do a fair amount of custom development um if you want to uh you know create these data as a product something that is very easy for your consumers to use and it's a pleasure to use you're going

(48:48) to have to build a lot of the tooling yourself so we for example in some of our clients we've we've had to write a lot of clis to be able to interact with the cloud services pull things together make it presentable in a decent sort of way because that just doesn't exist and I think as you know things are changing and the tooling is maturing but if you do embark on this this endeavor sort of at the moment you do need to be prepared for a fair amount of a fair amount of work so to sum up the key takeaways here

(49:20) um it's not really an All or Nothing exercise you can make progress as you move towards sort of each principle and you can still derive value as you as you do that it may not be sort of fully realized but it's it is something that you can do in terms of top tips make sure you start with the cloud it would it makes life a lot easier it's not to preclude a hybrid setup but it's in fact many organizations do but starting with the cloud as the main basis is is good for that be pragmatic about decentralization

(49:52) um you know you've got to sort of take these these things into account start small learn adapt rinse repeat and be prepared to do some custom integration tooling and development yourself so with that um just to conclude uh this is what we set out to answer in the beginning and we looked at that we said okay the considerations that we need to do if we're going to go on this journey is make sure we understand the reality of the situation we're in make sure we can ask the right questions and then choose

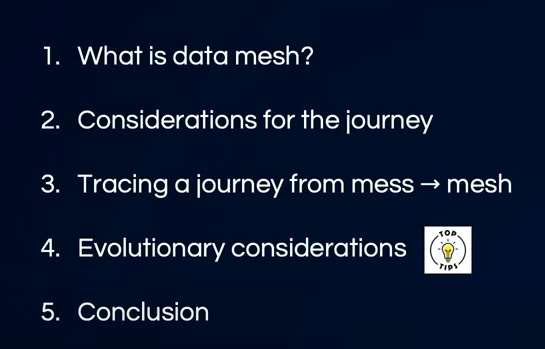
(50:22) the right the right approach for the job that may or may not be a data mesh but if it is we need to recognize that data meshes come in many shapes and sizes so what you do for one client or what you see in one organization is not directly transferable to another so data mesh is not the goal in and of itself it does provide a path for you to and good principles that will help you sort of be guided towards achieving your your objectives so we look through the example journey and we saw that you know even if datamesh is something that you're going

(50:58) to go for make sure you're prepared to go on the journey because there's a lot of work that's going to be required and if it's not recognized up front it can actually be a lot more problematic and finally The evolutionary sort of considerations uh that we saw it's not an All or Nothing exercise you can make progress along each principle and derive value and we've just gone through the the top tips and uh with that um that is the end so I I hope I hope that was helpful thank you thank you

We help clients to adapt and adopt emerging Technologies to solve their business problems

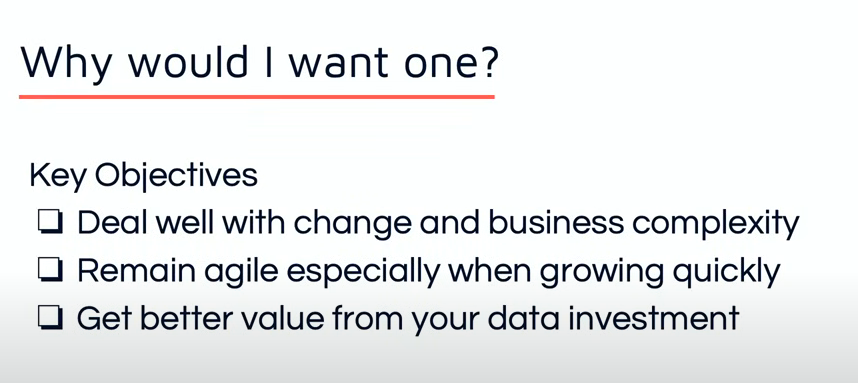
what I really hope is that I can give you some sort of insight into some of the things that we're seeing happening in this space

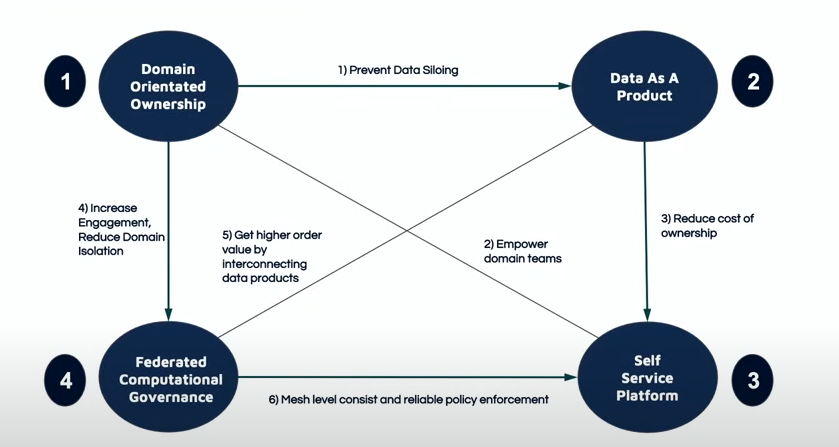
most organizations don't choose data mesh because all is going well they're usually unhappy with the status quo and they um they realize that they're in some kind of mess and they think that actually data mesh is potentially one of the answers to help us



It is new concept, new paradigm , way of operating created to address challenges faced by large oro trying to gain data driven insights from data at scale.

so it's a huge part of this is moving towards a more decentralized sort of model and approach





Domain orientated ownership what this really is about is it requires your organization generally to be broken up into domains so this is where people may have already done this if they've gone on like a microservice type transformation where you break the organization up into the areas

* rather than having for example your data being just dealt with by a central sort of Warehouse team
* instead of sort of having data just dumped into a central place where your analysts can go and get it and try and make sense of it
  + in a microservice architecture , you've got sort of cross-functional teams and stuff they will take responsibility for the services
  + Same for data, in addition to managing microservices ,the team and the domain now need to take responsibility for exposing what are called Data products out into the organization so that leads nicely on to the next
  + it it's now the responsibility of the individual domains and data product owner to expose what they have.

it's really treated with care so if you are exposing a data product

* you need to make sure that the the API is really stable
* you need to make sure that the there's a good quality you're not just putting some sort of unclean data
* you've got to make sure that the metadata exists so that people that are using the data can really understand
* the next principle is the self-service platform so this is really there to empower the teams and this is both the producers of the data product as well as those consuming it to be able to more easily build deploy and actually consume these data products

so the self-service platform is really there to enable the uh the technical side

* so if each team has to do all of this work themselves it's often too much and the the sort of cost becomes uh the burden becomes

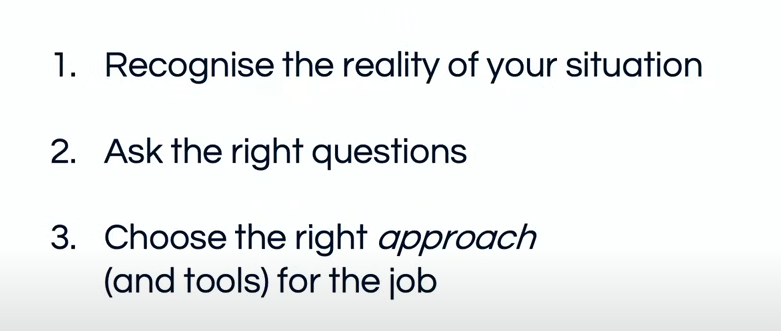
final one is Federated computational governance

* ownership is a movement of responsibility for the governance of the data to move down into the teams so rather than having a um a broad sort of governance group that decides everything the people in the team who know or in the domain who really know the the data best are the ones that are going to decide what is pii data what is sensitive data what is confidential data so that moves down into the into the domain's responsibility and you still need some level of oversight so hence the Federated part

the decentralization of this is really what helps you scale out because you don't need to coordinate as many things

**Considerations**

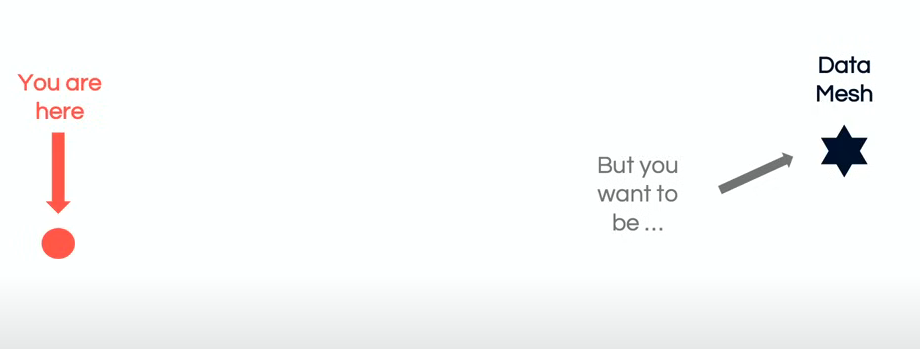
clients going on when they when they start to do this now often this will involve a fair amount of assumptions which are not necessarily correct and if it's not recognized or addressed early you may land up with more of a data mess than a mesh

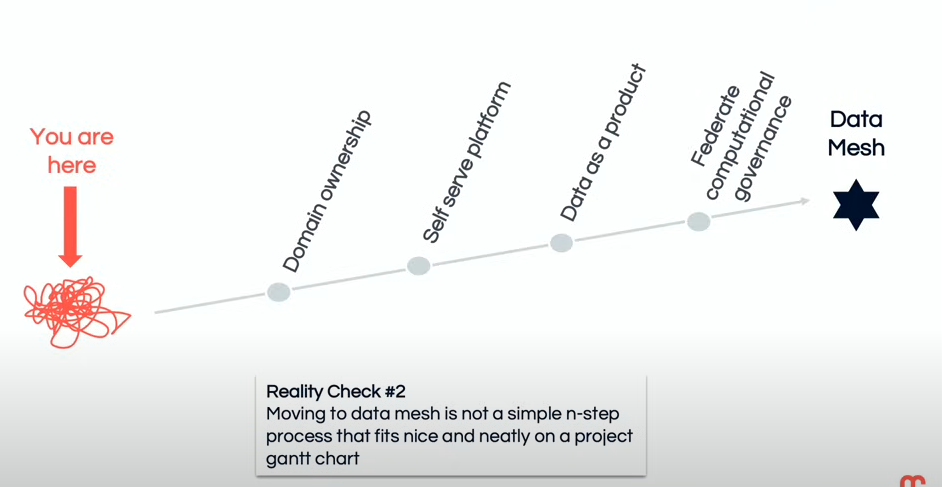


sort of anything, you need to have an honest look at your situation and ensure that you recognize the reality of it and this is understanding where you're starting from.

* oftentimes people make and organizations will make assumptions which turn out to be wrong and if we base our whole road map on that , it's it can be a real recipe for disaster
* I've got a problem uh data mesh is maybe the solution to help me get there how do I do that
  + we've read the book and we know there's four principles so we go through the principles
  + the second reality check kicks in and that is that moving to a data mesh is really not a simple sort of InStep process
  + in fact it's not actually steps at all it's more like a set of guard rails that try to help guide you in thinking about how to change your organization and your technical setup
  + we need to change our mindsets a little bit it's more going to be like a journey
  + we're gonna have to learn a few things go back on ourselves and as long as we do that we reckon we'll we'll get there except the third reality check will kick in which is that actually moving to data mesh is not just a journey that happens along a single a single path rather it's it's multiple Journeys that are all kind of interconnected
  + these are not standard lines um and this is to really indicate that these Journeys will really vary between different organizations so there's not a single way to implement data mesh
  + it's really going to be dependent on the context that you that you're in now

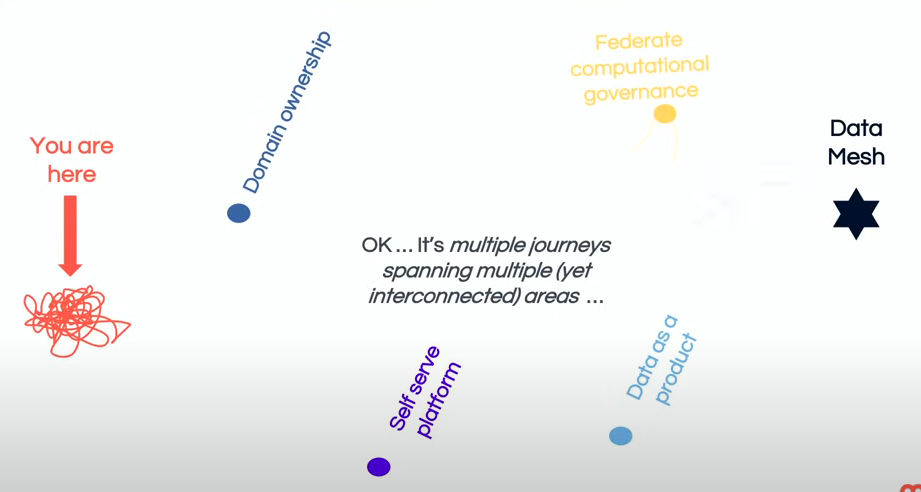
we need to change our mindsets a little bit it's more going to be like a journey

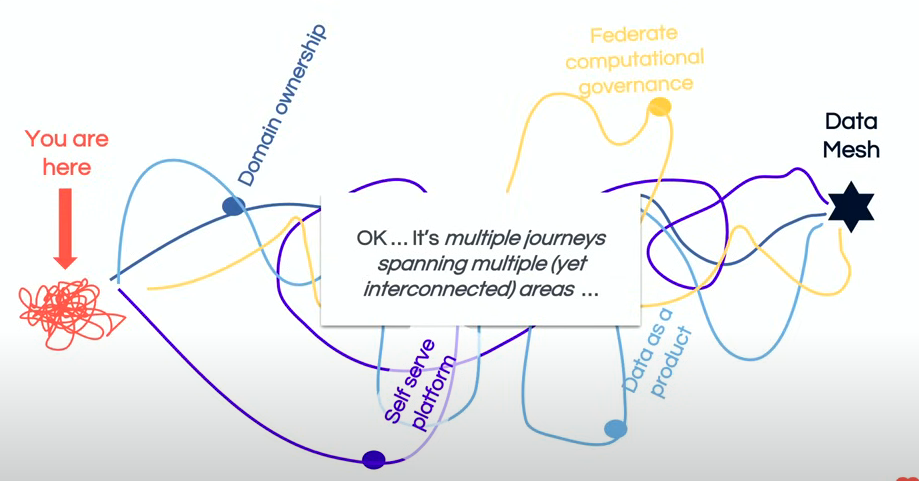












they're really

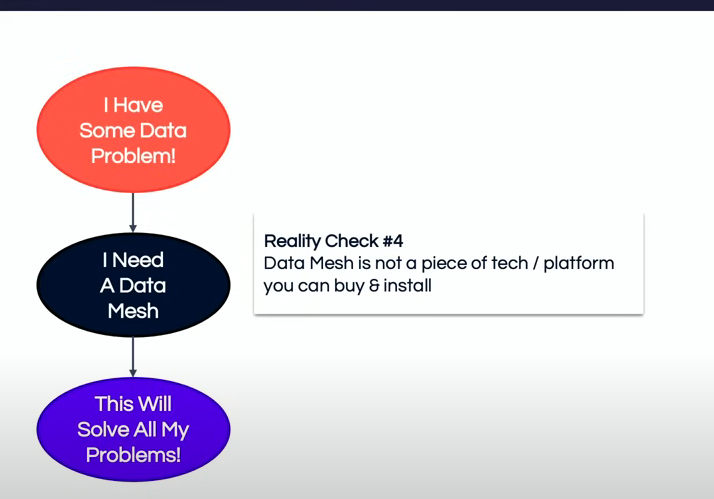
* interconnected so when you start doing something, it's more a little bit like this, so you kind of sort of start doing that, but it impacts some of the other areas. and as you as you move those as you move more towards, self-service platform , that will impact the other ones and you might not actually even get all the way there. but as you start doing this it does stabilize a little bit
* the key thing to realize here is that they all kind of interact with each other and it's not something that you're going to be able to just do one thing



* we should be able to get our data mesh and I'd say in terms of recognizing that it is sort of type of Journey that's required that's probably accurate except for one problem; you assume that you know that the answer is data mesh but how do you know that's true and how do you know that it will actually solve your problems

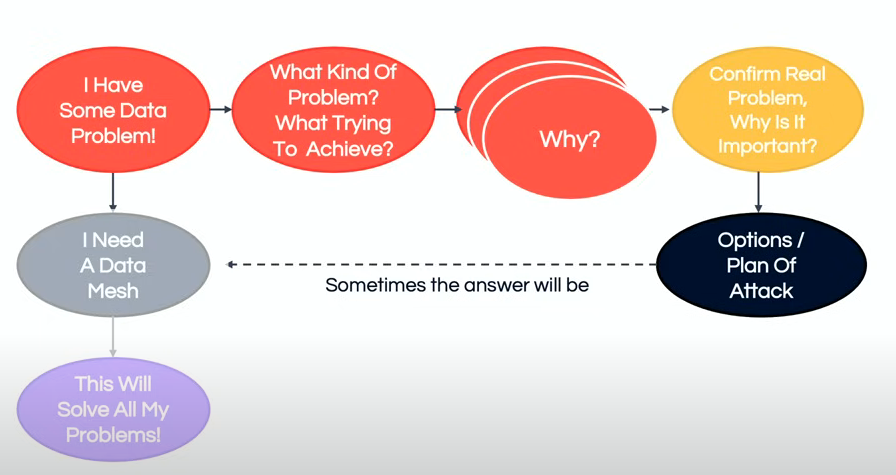
Reality check #4: Data mesh is not piece of tech/platform that you can buy/install

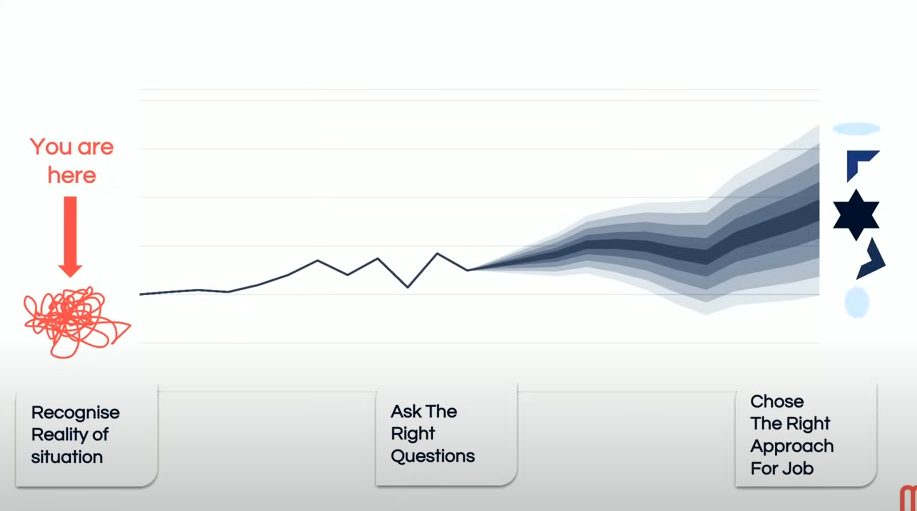
* company: I have some data problems:
* I need technology X (Data mesh)
* This will solve all my problem
  + It does not work like that: I said it's a socio-technical sort of approach that
    - requires both a change in your organization
    - your operating model
    - as well as the technical side of things
  + it's not something that you can just buy and kind of install

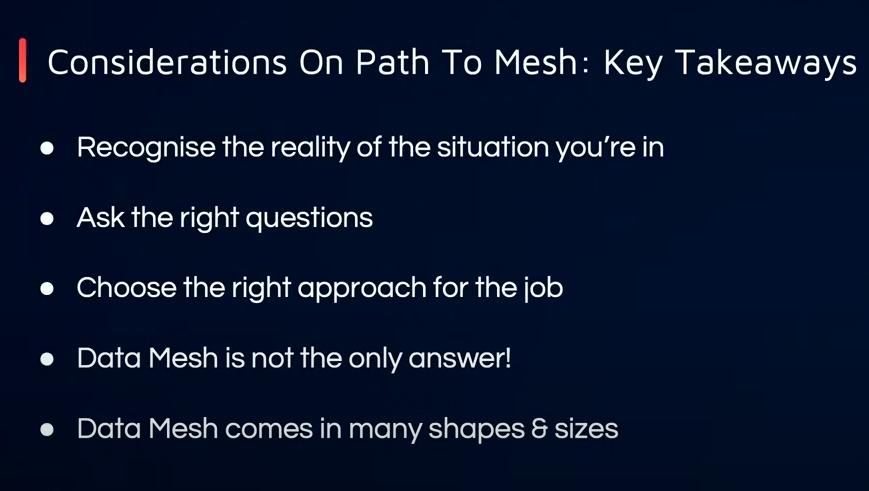


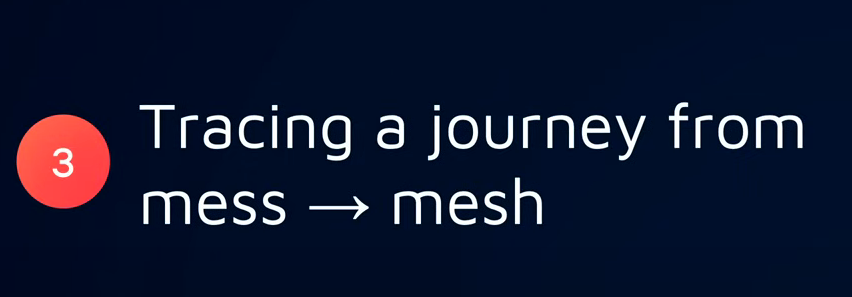
* the answer is going to be that that a data mesh is probably the right solution for you but sometimes it's not however, this is where it's really important to ask a second set of questions which is really about you know how far are you prepared to go in implementing this because you want to ensure you don't go down any sort of rabbit holes and waste time and money along the way
* you need to establish how close am I at the moment to being able to be in a position where I can start implementing it so it may be that you actually need to go through a digital transformation
* you need to incorporate and kind of put the data mesh objectives into that process and make sure that the requirements around that can be understood











however, there is no one-size-fits all so just because you've seen company a do it like this it doesn't mean that it's going to look the same for your company because it is completely context dependent

considering the options laid out to make this a little bit more practical let's look at some sort of real sort of problems that we see clients coming to us with and what they present with and see how how this works if we go through this process



* I'm not able to find my data easily there's too many systems that are all over the show
* I don't know where to start
* my starting point is a is a data Lake and so I just can't easily access or make sense of the data in there there's just this dump of files and things and as a an analyst I'm really struggling to be able to work out because with a data Lake architecture the files are usually dumped and the onus is on the consumer to be able to make sense